1.0 INTRODUCTION

Section 313 of the Emergency Planning and Community Right-to-Know Act (EPCRA) required the U.S. Environmental Protection Agency (EPA) to collect information concerning releases and other waste management activities of toxic chemicals to the environment from manufacturers, processors, and users of listed toxic chemicals. In order to collect such information, EPA implements a yearly reporting requirement from such facilities. Reports referred to as Form R chemical reports are due to EPA by July 1 each year to fulfill the reporting requirement for the previous calendar year. The reporting requirement was first implemented for the 1987 calendar year. The study discussed in this report reviewed data from the 1994 and 1995 reporting years (RY 1994 and RY 1995, respectively). Data for Standard Industrial Classification (SIC) Codes 25, 281, 285, and 30 were reviewed for RY 1994 and data for SIC Codes 286 and 26 were reviewed for RY 1995. SIC Codes 25, 281, 285, 30, 286, and 26 represent furniture manufacturing, inorganic chemical manufacturing, paint manufacturing, rubber and plastics manufacturing, organic chemical manufacturing, and pulp and paper manufacturing, respectively. This report also compares data for RY 1994 and RY 1995 to data from similar studies completed for the 1987 and 1988 reporting years. The data from the Form R chemical reports are compiled in EPA's Toxics Release Inventory database (TRIS) for future analysis, distribution, and evaluation. The information collected under EPCRA, Section 313 is useful for informing the general public and the communities surrounding affected facilities of releases and other waste management activities of toxic chemicals, assisting in focusing the Agency's research into the effects and control of toxic substances, and aiding in the development of regulations, guidelines, and standards.

For 1994, a total of approximately 76,500 Form R reports covering all SIC Codes required to report toxic chemicals were submitted to EPA by approximately 23,000 facilities and entered into the TRIS database. At the time the site surveys for RY 1994 were conducted, 12,896 Form Rs had been submitted and incorporated in the TRIS database for 3,764 facilities in SIC Codes 25, 281, 285, and 30. For RY 1995, a total of approximately 74,500 Form R reports covering all SIC Code codes required to report toxic chemicals were submitted to EPA by approximately 22,000 facilities. At the time the site surveys for RY 1995 were conducted, 402

facilities had been identified in the TRIS database as SIC Code 286 submitters and 165 facilities had been identified as SIC Code 26 submitters.

1.1 EPA's Overall Quality Assurance Program

Because of the wide audience and many intended uses of the Toxics Release Inventory database, EPA designed and implemented a program to assess the quality of the data collected under Section 313 and to identify areas where improved guidance would be useful for improving the accuracy of future reported data. The site surveys described in this report are a component of EPA's overall quality assurance program.

1.2 Site Survey Objectives

EPA's site surveys were designed to provide a quantitative assessment of the accuracy of the data submitted for a calendar year by identifying the frequency and the magnitude of errors in the Form R data and the reasons these errors occurred. EPA believed that on-site review of industrial processes, pollution control technologies, and documentation supporting the Form R reports would reveal errors in the database not obvious from review of a facility's Form R submissions. Expected error types included overlooked chemicals, incorrectly included chemicals, and errors in the release and other waste management quantity estimate calculations. The goal of the surveys was to obtain information that could be used to improve the Form R reporting instructions and definitions, and thus improve the quality of data in the TRIS database in future years.

Users of the results of the site survey program, as well as the TRI database itself, should be aware of a basic limitation of the Emergency Planning and Community Right-to-Know Act (EPCRA) reporting process. Under EPCRA (Title III of the Superfund Amendments and Reauthorization Act), facilities are not required to perform any additional monitoring or measurement of the quantity of toxic chemicals released to the environment to calculate Form R release estimates. Therefore, the methods selected by facilities to estimate releases and other waste management quantities depend on the nature of the data available to facility personnel, and the quality of these release and other waste management quantity estimates in turn depends on both the proper application of the estimation methods and on the quality of available data. At

facilities where supporting data were available, surveyors carefully examined the estimation calculations and data sources and then recalculated the estimates. In many instances, the site surveyors were able to identify data sources overlooked by facility personnel, and these new data were used to recalculate release and other waste management quantity estimates during the site visits. However, site surveyors did not conduct any monitoring or measurements during the site visits. Site surveyors also assessed the quality of the estimation methods by recalculating releases and other waste management quantities using alternative approaches where more accurate estimation methods were appropriate and where available data warranted.

1.3 EPA Site Surveys

EPA has conducted four sets of quality assurance site surveys since the first submittal of Form Rs from industry. The RY 1987 site surveys covered all SIC Codes affected by the EPCRA Section 313 (SARA Title III) requirements. The RY 1988 site surveys covered SIC Codes 28, 291, and 34 - 38. These SIC Codes were targeted because data for the 1987 reporting year showed that facilities in these SIC Codes accounted for a substantial portion of the total releases from all reporting facilities in 1987.[1]

The approach used for the RY 1994 and RY 1995 site surveys was similar to that used for the RY 1987 and RY 1988 programs. Training of site surveyors, the contents of the survey instrument, and activities conducted on site for RY 1994 and RY 1995 were similar to the previous programs. However, the SIC Codes included in the site visits differed slightly from those studied in previous years. The RY 1994 site surveys focused on facilities in SIC Codes 25, 281, 285, and 30; and the RY 1995 site surveys focused on facilities in SIC Codes 26 and 286. These SIC Codes were targeted because previous reporting years showed that facilities in these SIC Codes account for a substantial portion of the total releases and other waste management quantities. The results of the RY 1994 and RY 1995 site surveys will help EPA identify ways additional guidance can be structured to improve the overall quality of the data generated under EPCRA (SARA Title III, Section 313) reporting.

---

[1]The results of these surveys are provided in Radian reports entitled Assessment of Data Quality in the 1987 Toxic Release Inventory: Site Visit Program (March 1990), and Site Visit Program to Assess 1988 Toxic Release Inventory Data Quality (July 1991).

2.0 APPROACH

A very structured approach was established for the site surveys to ensure consistency in conducting site surveys and accuracy of the results. The approach was orginially established for the RY 1987 and RY 1988 site surveys and was improved for the RY 1994 and RY 1995 site surveys based on experience from the previous programs.

The approach for performing the RY 1994 and RY 1995 site surveys, shown schematically in Figure 2-1, consisted of the following steps:

(1) Revising the Survey Instrument;

(2) Selecting facilities to be visited (Sample Selection);

(3) Training site surveyors (Training);

(4) Arranging Site Visits;

(5) Performing site visits (Site Visit Methodology);

(6) Data Management and Data Quality Assurance; and

(7) Data analysis and Reporting.

Each of these steps is discussed in the following subsections.

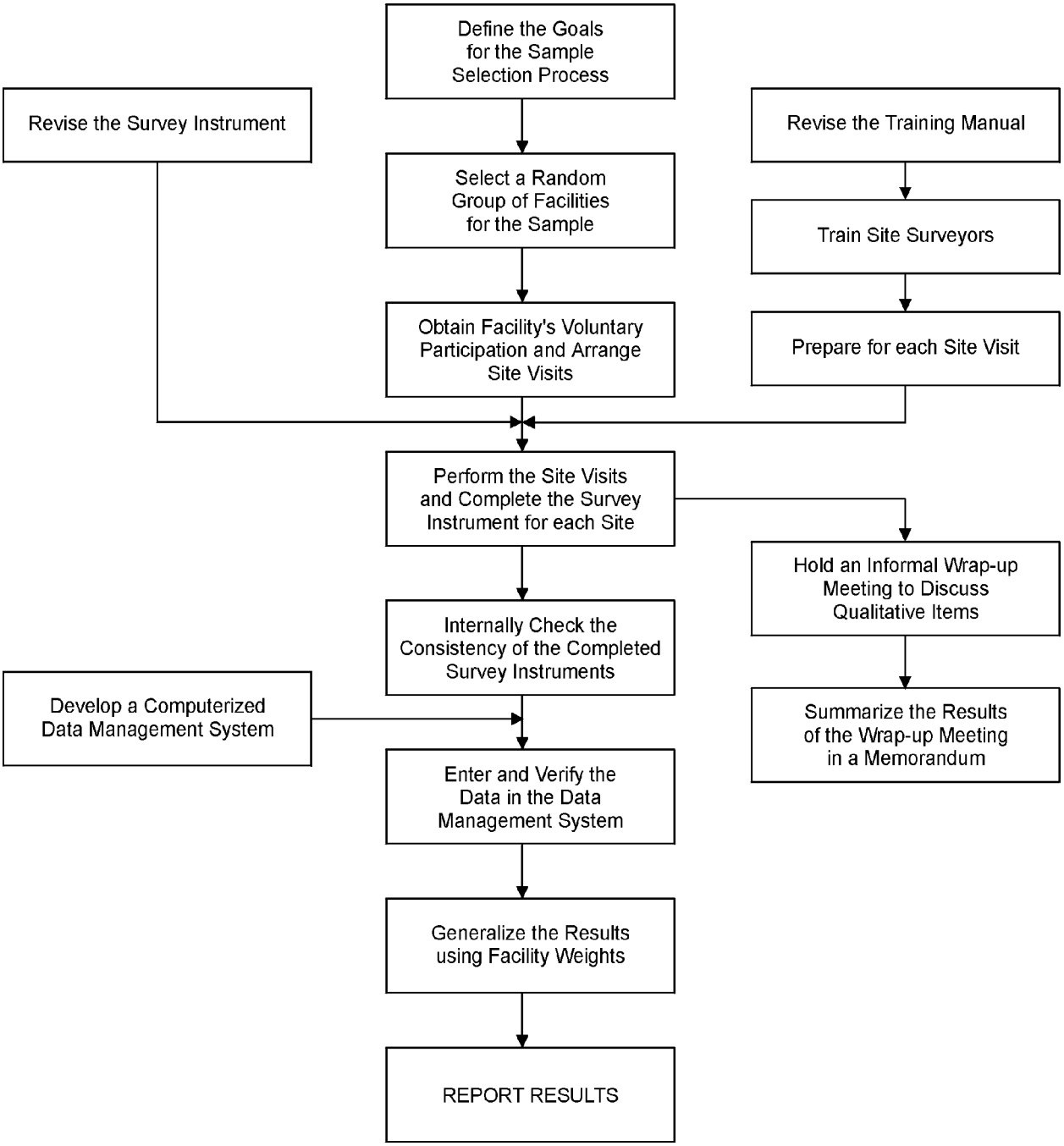

Figure 2-1. Approach used to Perform the EPCRA Section 313 Site Visit Program

2.1 Survey Instrument

The survey instrument, shown in Appendix A, was designed to standardize and facilitate the review of threshold determinations, release estimate calculations, and calculations used to assess other waste management activities at facilities. The engineers and scientists who performed the site surveys used the survey instrument as a detailed checklist to ensure that all pertinent items were reviewed, and also as a consistent format for recording both the data collected during site surveys and the errors made by facility personnel on their Form R reports. In addition to its primary focus on chemical-specific information, the survey instrument contains questions regarding the usefulness of the reporting instructions, EPCRA Section 313 hotline, and the other published guidance materials. A question on the usefulness of the Toxic Chemical Release Inventory Form A (Form A) was added for the visits conducted for RY 1995. Each section of the survey instrument focuses on identifying specific types of errors made by facility personnel on their Form R submittals.

The survey instrument used in the RY 1994 and RY 1995 data quality site visits was a revised version of the survey instrument used in the RY 1987 and RY 1988 programs. Most of the questions remained the same, but some additional questions concerning documentation available, possible sources for threshold determinations, source reduction activities, pollution prevention technology, and use of the Form A (for RY 1995) were added to clarify information received during the site visits and to assess the usefulness of the new guidance and materials available. The time increments for amount of time needed to complete all Form Rs at the facility were adjusted slightly in the RY 1995 survey instrument to obtain a more precise estimate of time needed. The format was also revised to make the survey instrument easier for the site surveyors to use.

2.2 Sample Selection

The primary objective of sample selection was to obtain a random group of facilities from the key industry groups within specific SIC Codes which were being surveyed to appropriately scale up the results to reflect the reporting of the entire SIC Code group. This

sample selection approach was used for Reporting Year 1987, 1988, 1994, and 1995 survey programs.

2.2.1 RY 1987 and RY 1988 Sample Selection

All industry groups required to report toxic chemicals under the EPCRA Section 313 program were first surveyed for Reporting Year (RY) 1987. Table 2-1 presents the distribution of facilities sampled among the SIC Codes for each year of the data quality site visits.

A target of 150 facilities was selected as the number of facilities needed to ensure the statistical validity of the data collected during the site visit program for RY 1987. Appendix B provides a detailed discussion of the procedure used to select the sample group of facilities, and also provides a description of the weighting system (i.e., the number of facilities in the TRI database represented by each sample point). Briefly, facilities submitting 15 or fewer Form R reports were divided into geographic clusters on the basis of the first three digits in their zip codes. A sample of geographic clusters was then selected according to a sampling scheme in which probability of selection is proportional to cluster size. The cluster size measure was the total number of facilities in the SIC Code group sampled. The clustering approach was used to minimize costs by reducing travel costs and travel time for site survey teams. A stratified random sample of facilities was drawn from each of the sample geographic clusters, based on the desired number of site visits in each SIC Code. This general procedure was used for sample selection for site surveys conducted for RY 1987 and RY 1988.

Facilities with 15 or fewer Form Rs were selected due to the limited time and budget available. Only a few facilities have 16 or more Form Rs and site visits to those facilities would have taken considerable time, limiting the number of facilities that could be visited. Since the same facility personnel may complete multiple reports at a given facility, visiting more facilities presents a better representation of the range of reporting practices.

Table 2-1

Distribution of Facilities Among the SIC Codes For Each Year of the Data Quality Site Visits

SIC Code	Number of Facilities Visited			
	RY 1987	RY 1988	RY 1994	RY 1995
20	16			
22	5			
23	1			
24	2			
25	2		25	
26	14			10
27	3			
28	44	43	37[a]	10
29	0	1		
30	7		23	
32	2			
33	16			
34	16	8		
35	5	10		
36	11	14		
37	7	10		
38	2	3		
39	3			
Total	156	89	85	20

[a]One of the facilities visited was in SIC Code 282. The results of this survey are not included in the analysis of data for SIC Code 281 and 285.

A target sample size of 90 completed site visits was the goal of the site visit program for RY 1988. Details of the sample design and weighting methodology are described in Appendix B (and follow the same general procedure as RY 1987). For RY 1988 facilities submitting 30 or fewer Form R reports were targeted, rather than facilities with 15 or fewer Form Rs as in other years.

2.2.2 RY 1994 and RY 1995 Sample Selection

The key industries sampled for the RY 1994 Toxic Release Inventory (TRI) data quality site survey were furniture manufacturing, paint manufacturing, chemical manufacturing, and rubber and plastics manufacturing. Key industries sampled for the RY 1995 TRI data quality site survey were the organic chemicals manufacturing and pulp and paper industries. Facilities in these industry groups were selected because they accounted for a substantial portion (approximately half) of the total releases reported by facilities for the 1994 and 1995 reporting years. The sample does not include facilities outside the above listed industry groups and therefore does not represent the entire population of facilities that reported to the TRI.

Facilities engaged in furniture manufacturing were defined as those having a two-digit SIC Code of 25. Facilities engaged in chemical manufacturing (SIC Code 28) were ultimately refined to include only those facilities engaged in inorganic chemical manufacturing with an SIC Code of 281, and paint manufacturing with an SIC Code of 285. Rubber and plastics manufacturing facilities have an SIC Code of 30.

A target sample size of 40 completed site visits, divided evenly between SIC Codes 281 and 285, was established for the first part of the RY 1994 site visit program. A target sample size of 50 completed visits, divided evenly between SIC Codes 25 and 30, was the goal of the second part of the RY 1994 site visit program. As discussed in Section 2.2.1, a stratified random sample of facilities was drawn from a set of geographic clusters. The 1995 site visit program targeted 20 completed site visits at facilities in SIC Codes 286 and 26 (10 visits each). The geographic clustering approach was not used for RY 1995 because the sample set in the SIC Codes chosen was small. A total random sampling was done for RY 1995. Details of the sample design and weighting methodology are described in Appendix B.

2.3 Site Surveyor Selection and Training

To complete the site visit program as efficiently as possible, the engineers and scientists staffing the program were selected on the basis of their experience in performing environmental audits of industrial processing facilities, and were required to have a thorough understanding of chemistry, technical calculations, multimedia environmental concerns, and pollution control technologies. The quality assurance reviewers for the site surveys were all from one office and remained consistent throughout the program.

A surveyor training program was developed to ensure consistency and high quality work among all site surveyors. The training program consisted of three steps:

1) Compiling a comprehensive training manual, including copies of EPA guidance documents and other references;

2) Holding training sessions to familiarize project personnel with program requirements; and

3) Review of the completed survey instruments with the site surveyor by the reviewer to maintain a consistent approach among the surveyors.

2.4 Arranging Site Visits

The goal in arranging site visits was to provide each facility in the sample with an equal opportunity to participate in the site visit program, thus ensuring the statistical validity of the approach. Participation was voluntary; the facilities were not legally required to participate. A key factor encouraging voluntary participation was the assurance of anonymity to the facilities. Names, location, and all other facility identification data are shielded from the Agency. Upon facility request, a written confidentiality agreement was signed by the contractors.

As a first step, introduction letters (copies of these letters are provided in Appendix C) were sent to each facility's technical contact, and where appropriate, to each facility's senior management official. These letters contained explanations of the purpose of the quality assessment program and the anticipated burden on and benefits to the facility, and

assurance to the facility that all facility-specific data would be treated as confidential. ERG followed these letters with telephone calls to the technical contacts at the facilities to solicit their participation, and for those facilities agreeing to participate, to arrange a date for the site visit and to review a preliminary agenda for the site visit.

2.5 Conducting Site Visits

The goal of the site visit was to collect all the information needed to complete the survey instrument accurately, while minimizing burdens on facility staff. On-site survey activities included tours of the facilities, which focused on material storage areas, industrial processing operations, and pollution control equipment; careful review of all readily available documentation, which could include MSDSs, production data, monitoring data, purchasing records, and facility spreadsheets or computer software with this information; and interviews with appropriate facility employees regarding documentation materials. Site surveyors did not perform any monitoring or measurements during the site visits.

The site visits were designed to determine:

1) Overlooked chemicals;
2) Releases and other waste management activities;
3) Errors in the Form R reports submitted to EPA; and
4) Whether more accurate release estimation methods could have been used, based on information available to the facilities.

Releases and other waste management estimates were either recalculated or recreated by site surveyors from available documentation during the visit. Site surveyors recorded these results on the survey instrument and reviewed the results with facility personnel before leaving the site. A wrap-up meeting at the facility with the person who filled out the Form R reports was held at the end of the visit to discuss any issues or questions that the facility contact had and to go over the conclusions and recommendations of the site surveyor. Follow up with the facility contact after the on-site visit occurred when regulatory issues which needed EPA clarification or additional research was required.

2.5.1 Data Collection

Site surveyors reviewed 295 Form R chemical reports and 728 additional chemicals with amounts used or activities which did not meet the reporting criteria at the 85 facilities visited for RY 1994, and 139 Form R chemical reports and 171 additional chemicals with amounts used for activities which did not meet the reporting criteria at the 20 facilities visited for RY 1995. Threshold determinations, releases, and other waste management estimates were reviewed separately to identify the frequency, magnitude, and sources of errors in these areas. Site surveyors followed the stepwise approach described in the Form R reporting instructions for completing threshold determinations, releases, and other waste management estimates. In following the Form R reporting instructions, facilities must first assess which chemicals are manufactured, processed, or otherwise used in excess of appropriate thresholds. Facilities must then estimate and report all releases to the environment and other quantities of listed chemicals exceeding a threshold managed as waste.

2.5.2 Threshold Determinations

The following types of errors may be made by facilities in determining which chemicals at their site meet a EPCRA Section 313 thresholds:

- Overlooking a chemical;
- Incorrectly calculating a threshold amount;
- Incorrectly applying an exemption; and
- Misclassifying a chemical activity.

To identify errors in threshold determinations, site surveyors looked for problems in a facility's documentation and, on the plant tour, site surveyors looked for evidence of chemicals that were reported but should not have been reported, and for evidence of chemicals that were not reported but should have been reported. Each facility's documentation was reviewed to track the decision process used to determine whether a chemical should have been reported. Furthermore, site surveyors used all available documentation to recalculate threshold

estimates for reported chemicals and for chemicals present but not reported to verify the accuracy of facility calculations.

2.5.3 Release and Other Waste Management Estimates

The following types of errors may be made by facilities in calculating release estimates for EPCRA Section 313 chemicals:

- Overlooking a chemical;
- Overlooking a source of data;
- Incorrectly calculating a release or other waste management quantity; and
- Incorrectly interpreting the reporting instructions.

A two-part approach was used for identifying errors in releases and other waste management activity estimates. First, site surveyors always recalculated releases and other waste management quantities using the same technical approach used by the facility. Second, whenever the site surveyor's experience and training indicated that a calculation approach different than that used by the facility was appropriate, the surveyor attempted to obtain the data needed to calculate releases and other waste management quantities using the more appropriate approach. In many such instances, data were not readily available during the site visit to recalculate these amounts using the alternative approach. In the cases where site surveyors were able to recalculate releases and other waste management amounts using alternative approaches, they were able to assess the reasonableness of the estimation techniques used by facility personnel.

The surveyors quantified all numerical differences between the facility's estimates and the recalculated values, even in instances where surveyors identified only small differences. As discussed later, these numerical differences were used to assess quantitatively the accuracy of the total aggregate releases and other waste management quantities contained in the TRI database.

2.6 Data Management/Data Quality Assurance

Many steps were taken to ensure the data quality of the surveyor's estimates and the verification of the data in the database. This section outlines the procedures taken to review the Survey Instruments after they have been completed by the site surveyor, the database system, and the data entry into the master database; the verification procedures for the data entered into the database; the weighting of the data to apply the results to the entire population of facilities for each SIC Code surveyed; and potential sources of error in the site survey program.

2.6.1 Quality Review of Survey Instrument and Data Entry

All survey instruments were reviewed twice by a consistent set of reviewers to ensure the calculations and methodologies used were correct and consistent for all site surveys. The data entry for all site surveys was also done twice. These database entries were compared to each other, and then verified with the actual survey if an inconsistency was found. Project staff also reviewed the database entries for internal consistency and completeness by comparing responses to various questions as appropriate.

2.6.2 Data Weighting

To allow EPA to assess the impact of the site survey program results on the TRI database for the SIC Codes surveyed, weighting factors were applied to the site visit data. These factors or "weights" represent the number of facilities in the TRI database represented by each of the surveyed facilities. The weights of each surveyed facility are based on the measure of size of the geographic cluster in which the facility is located, and the systematic probability of selecting that facility proportional to that measure of size.

The weights for the sample facilities in each SIC Code group are summed up to represent the total population of facilities included on the TRI for that SIC Code group. A total population of 535 facilities for SIC Code 25, 1872 facilities for SIC Code 30, 662 facilities for SIC Code 285, and 695 facilities for SIC Code 281 is represented for RY 1994. A population of

402 facilities for SIC Code 286 and 165 facilities for SIC Code 26 is represented for RY 1995. The weights used for facilities in SIC Codes 25, 281, 285, and 30 are presented in Appendix B.

2.6.3 Limitations of the Analysis

The design and implementation of the survey may have introduced unavoidable inaccuracies in the study results. The three primary sources of error are:

- sample selection bias;
- survey implementation; and
- data reduction and analysis.

The relatively small number of facilities sampled clearly introduced a sample selection bias -- the smaller the number of facilities sampled the greater the likelihood that these facilities do not accurately represent the universe of reporting facilities. For the selected sample size of 40 facilities in SIC Code 28, the 90 percent confidence interval is plus or minus thirteen percent. That is to say, if 50 percent of the facilities visited reported accurate data there is a 90 percent probability that between 37 percent and 63 percent of the facilities in the national database reported accurate data. Counting the SIC Code groups separately, there is a 90 percent confidence level of plus or minus 18 percent for each group. For the selected sample size of 50 facilities in SIC Codes 25 and 30, the 90 percent confidence interval is plus or minus 17 percent. Counting the SIC Code groups separately, there is a 90 percent confidence level of plus or minus 24 percent for each group. Thus, the confidence levels are based on the survey size and the total number of facilities in the SIC Code group.

Another possible source of error concerns the fact that approximately 15 different surveyors performed the survey. This source of inaccuracy was controlled to the extent possible by the use of a carefully designed survey instrument and extensive quality assurance provisions. Nevertheless, it is possible that different surveyors made different judgments in the course of the site surveys.

Finally, certain assumptions were made to simplify data analysis. The key assumption was that the facilities and Form Rs examined in the site visits accurately represent all facilities in their SIC Code group in terms of the accuracy of the data submitted. Aside from possible errors introduced by the relatively small size of the sample, the sampled facilities may not fully represent their SIC Code group because:

- The sampled facilities excluded any facility with more than 16 Form Rs for budgetary reasons. To the extent that facilities submitting more than 16 Form Rs report more (or less) accurate data than the sampled facilities, the latter facilities do not fully reflect the universe of facilities in the database.

- Many facilities surveyed processed or manufactured some kind of specialty chemical. These facilities may not accurately portray the "typical" facility within the SIC Code group. This may overestimate a specific chemical produced and released within the SIC Code group due to scaling and weighting factors.

2.7 Data Analysis and Reporting

Once the results of the site surveys were loaded into a database and the database was validated through the quality assurance process described above, the data were evaluated to discern trends in the quality of data in the TRI forms. This report presents the results of that analysis.